



Gouvernance informationnelle de l'environnement et partage en ligne des données publiques. Politiques et pratiques de l'opendata environnemental (Amérique du sud - France).

Pierre Gautreau, Matthieu Noucher

► To cite this version:

Pierre Gautreau, Matthieu Noucher. Gouvernance informationnelle de l'environnement et partage en ligne des données publiques. Politiques et pratiques de l'opendata environnemental (Amérique du sud - France).. NETCOM: Réseaux, communication et territoires / Networks and Communications Studies, 2013, 27 (1-2), p .5-21. 10.4000/netcom.1234 . halshs-00991878

HAL Id: halshs-00991878

<https://shs.hal.science/halshs-00991878>

Submitted on 25 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GOUVERNANCE INFORMATIONNELLE DE L'ENVIRONNEMENT ET PARTAGE EN LIGNE DES DONNEES PUBLIQUES

POLITIQUES ET PRATIQUES DE L'OPEN DATA ENVIRONNEMENTAL (AMERIQUE DU SUD - FRANCE)

GAUTREAU PIERRE¹, NOUCHER MATTHIEU²

Ce numéro thématique³ a pour ambition de contribuer à une meilleure compréhension des formes actuelles de gouvernance informationnelle de l'environnement en explorant le cas spécifique du partage en ligne des données des pouvoirs publics. La notion de « gouvernance informationnelle » de l'environnement désigne deux évolutions majeures marquant les formes de gestion des questions environnementales depuis les dernières décennies du XXe siècle : la constitution de l'information en un nouveau pilier des politiques publiques, aux côtés des instruments réglementaires et économiques (Lavoux, 2003), ainsi que le constat que l'information n'est pas qu'un simple matériau nécessaire à la formulation de politiques, mais une ressource possédant un potentiel de transformation des formes de gouvernance, qui « restructure les processus, les institutions et les pratiques » (Mol, 2009). L'émergence de cette gouvernance informationnelle n'est pas propre au champ environnemental, mais elle s'y manifeste par une grande diversité d'expressions : multiplication des instances participatives fondées sur la discussion ou la co-construction d'information pour le traitement d'un nombre croissant de problèmes ; généralisation des pratiques d'évaluation d'impact environnemental ; pilotage des politiques territoriales par la production d'indicateurs ; évaluation et comparaison des pratiques de gestion de l'environnement ; essor des normes par le biais de la certification, de la labellisation ou du rapportage extra-économique des entreprises (Alphandéry et al., 2012a). Selon Mol (2009), un des signes de cette importance désormais cruciale de l'information est le recentrage d'une grande part des actions des mouvements sociaux autour des

¹ Maître de Conférences, Université Paris 1 Panthéon Sorbonne, UMR PRODIG 8586. 2 rue Valette, 75005 Paris. Tél. : 0033 (0)1 44 07 75 99 pierre.gautreau@univ-paris1.fr

² Chargé de recherche au CNRS, UMR ADESS 5185 (CNRS / Universités de Bordeaux), Maison des Suds 12 esplanade des Antilles 33607 Pessac. Tél. : 0033 (0)5 56 84 82 06 matthieu.noucher@cnrs.fr

³ Le projet de ce numéro spécial est issu du Projet Baguala « Usages des données environnementales en accès libre en Amérique du Sud et en France » (CNRS-Université Paris 1 Panthéon Sorbonne, <http://baguala.hypotheses.org>).

questions environnementales autour de demandes d'accès à l'information, à côté des actions classiques visant à modifier législations et pratiques. Les évolutions technologiques constituent par ailleurs une des dimensions et une des conditions de l'émergence de cette gouvernance environnementale. Elles modifient simultanément les formes de construction des connaissances (numérisation des données, systèmes d'information) et les formes de leur circulation (diffusion, réutilisation), impactant donc sur la formulation des problèmes environnementaux.

L'essor d'initiatives de type *Open Data* est souvent perçu comme emblématique de cette révolution en cours, matérialisant à la fois la réponse des pouvoirs publics à la transformation des exigences démocratiques et les possibilités offertes par la technique au déploiement et à l'actualisation de ces aspirations sociales. Mettant au premier plan la notion de « partage » des données, l'Open Data est censé représenter une forme de renouvellement du pacte entre Etat et citoyens, autour de l'échange d'informations. Mais les contours et les modalités de ce renouvellement sont mal connus, et le sens même de cette ouverture est au cœur des controverses : « *moyen de pression pour la tendance libérale, elle témoigne d'une véritable méfiance envers toute institution publique pour les libéraux-libertaires tandis que d'autres acteurs y voient une opportunité de développer la démocratie participative* » (Giraud, 2014). Ainsi, ceux qui se présentent volontiers comme les militants du « mouvement open data » (Chignard, 2012) sont porteurs de valeurs, d'objectifs et de stratégies qui diffèrent de ceux de la puissance publique.

L'« open » d'Open Data désigne sans doute moins ce qui rassemble les différents acteurs de ce mouvement que ce qui les différencie⁴. A l'instar de Pierre-Amiel Giraud (*ibid*), nous préférons alors utiliser l'expression « *mouvance Open Data* », pour signaler que le terme d'Open Data regroupe des approches émergentes aux finalités fluctuantes et sans véritable cadre unificateur. C'est précisément parce qu'elle déborde de tout cadrage réglementaire ou discours militant et se traduit par une variété d'acteurs et de pratiques que la mouvance Open Data nous semble un objet de recherche intéressant pour mieux aborder la gouvernance informationnelle de l'environnement. Les différentes contributions de ce numéro thématique se focalisent en particulier sur certaines transformations du rôle de l'Etat dans un domaine où il a longtemps été un acteur dominant et incontesté. Les propositions ici réunies mobilisent des approches complémentaires, au-delà du regard géographique qui leur est commun : études de l'Etat, études des sciences et techniques, *Political ecology* des savoirs sur l'environnement.

L'article de Maya Leroy et Sandra Nicolle -*Stratégies de partage et diffusion de données publiques environnementales : cas d'étude en Amazonie française et brésilienne*- compare les pratiques des autorités publiques en terme de partage des données sur la déforestation et l'orpaillage clandestin au Brésil et en Guyane française. En replaçant ces pratiques dans leurs contextes nationaux respectifs, les auteures éclairent le rôle attribué à l'information dans les politiques des deux territoires : l'affirmation

⁴ C'est ce que note J. Chibois, dans un billet intitulé « Que signifie l'« open » de l'Open Data ? Du patch au hack, et de la publicisation à l'ouverture paru sur : <http://laspic.hypotheses.org/1896>

internationale d'une souveraineté sur les ressources amazoniennes dans le premier cas, un contrôle répressif des activités d'orpaillage dans le second. L'option du Brésil pour une ouverture affirmée de ses données de déforestation contraste avec le processus à rebours en Guyane française. Dans *Logiques de projet et régulation publique de l'information géographique: l'expérience bolivienne*, Louca Lerch interroge les fonctions attribuées au partage de l'information géographique par l'Etat bolivien tel qu'il évolue à partir de l'élection de Evo Morales en 2005. Cette étude de cas détaille une initiative publique de collecte et de mise à disposition de données géographiques numérisées relatives au territoire bolivien. Elle éclaire de façon originale le rôle que peut jouer l'information environnementale dans un projet politique de reconstitution des capacités étatiques de connaissance des ressources nationales, faisant suite au démantèlement de pans entiers de l'administration durant la période d'ajustement structurel des années 1990. Dans l'article *Formes et fonctions de la « donnée » dans trois webs environnementaux sud-américains (Argentine, Bolivie, Brésil)*, Pierre Gautreau, Marta Severo, Timothée Giraud et Matthieu Noucher analysent la diversité des pratiques de mise en ligne de données environnementales, remplaçant celles de l'administration publique au sein du continuum de l'ensemble des acteurs du web. Cette approche débouche sur un triple constat : l'intérêt de l'étude du web pour repérer les acteurs les plus actifs dans la mobilisation de l'information environnementale au service de leurs stratégies politiques, le rôle paradoxalement marginal de la fourniture de données dans la hiérarchisation des acteurs du web environnemental, la mise en évidence des stratégies spécifiques de l'Etat pour se maintenir comme un acteur central du web environnemental dans un contexte de forte concurrence des acteurs privés et associatifs. Enfin, l'article de Matthieu Noucher *Infrastructures de données géographiques et flux d'information environnementale : de l'outil à l'objet de recherche* esquisse le potentiel heuristique de l'étude d'institutions nées dans les années 2000 et ayant fait jusqu'alors l'objet de peu d'investigations en sciences sociales. Les IDG ou Infrastructures de Données Géographiques sont des portails Web dont se dotent un nombre croissant de territoires. Elles renouvellent en profondeur les logiques de partage des connaissances cartographiques et de construction des savoirs institutionnels sur le territoire et l'environnement. Leur caractérisation comme des dispositifs socio-techniques d'un genre nouveau ouvre des pistes de recherche originales pour mieux comprendre comment les pouvoirs publics influencent l'évolution des formes de circulation des informations environnementales.

Ces articles sont complétés par une série de 14 notices qui présentent de manière synthétique les contours techniques, politiques et organisationnels d'initiatives variées⁵. Ces retours d'expériences concernent ainsi aussi bien des projets internationaux -comme le Global Biodiversity Information Facility (GBIF) ou l'Inter-American Institute Data and Information System-, nationaux -comme l'Inventaire National du Patrimoine Naturel en France ou le système d'information sur la

⁵ On trouvera une autre série d'exemples d'initiatives concernant essentiellement des pays latino-américains, sur le site du colloque de Porto Alegre d'août 2012 « Sharing environmental information: issues of open environmental data in latin america » (<http://baguala.hypotheses.org/publications/seminar-porto-alegre-2012>).

biodiversité de l'administration des Parcs nationaux argentins-, régionaux -comme le portail des données naturalistes SILENE en région PACA ou l'infrastructure de données géographiques GéoGuyane- ou locaux, comme la plateforme PIGO de Ouagadougou. Ces notes d'information illustrent aussi la variété des acteurs qui portent ces initiatives : centres de recherche (pour le réseau qui diffuse des données sur les capteurs de température des eaux côtières dans la région du Pacifique Sud et Sud-Ouest), services de l'Etat (pour l'application de la Directive Inspire), associations (Ligue de Protection des Oiseaux, Tela Botanica), collectivités territoriales (avec le cas de la Ville de Montpellier). Enfin, l'entretien avec Simon Chignard, auteur du livre « L'open data, comprendre l'ouverture des données publiques », paru en 2012, permet de compléter ce panorama en s'interrogeant sur les spécificités des données environnementales au sein des projets Open Data français qui tendent à se généraliser.

Cet ensemble de textes (articles, notes d'information, entretien) couvre à la fois le contexte français et sud-américain, éclairant des enjeux géographiquement différenciés, certains aspects des débats portant sur l'Open Data environnemental et permettant de mieux cerner quelques difficultés d'étude de cet objet complexe.

1. DES DIFFICULTES A CERNER CE QUE RECOUVRE LA MOUVANCE DE L'OPEN DATA (ENVIRONNEMENTALE)

La mouvance de l'Open Data échappe à une caractérisation aisée principalement pour deux raisons, relatives au sens à attribuer à ces initiatives dans l'éventail des actions de l'Etat. Tout d'abord, les motifs du partage des données sont généralement multiples, flous et enchevêtrés. Ensuite, bien que l'Open Data soit présenté comme un outil parmi d'autres de l'action publique, nombre de ses caractéristiques en font un objet à part des « instruments » les mieux établis tels que « le recensement, la cartographie, la réglementation, la taxation... » (Lascoumes et Le Galès, 2004a), rendant malaisée la saisie de ses effets sociaux et territoriaux.

1.1. Les raisons du partage des données publiques : entre optimisation managériale et exigences démocratiques

Pour rendre compte de ce qui pousse une administration publique à partager ses données environnementales, plusieurs explications alternatives peuvent être avancées. Selon les contextes, ces motifs peuvent s'exclure, se juxtaposer, voire se renforcer mutuellement. L'explication générale qui nous semble rendre le mieux compte de l'essor de ces initiatives consiste à les interpréter comme la mise en pratique de théories développées dans le cadre du paradigme de la « société de l'information » et d'une diffusion à l'administration publique de formes entrepreneuriales de gestion relevant du « nouveau management public »⁶. Dans ce

⁶ Hood C. 1995. Contemporary public management: a new paradigm ? *Public policy and administration* 10(2).

cadre, l'information, avant d'être un droit social à défendre, est une ressource pour l'optimisation de la gestion environnementale, un outil de plus pour la modernisation écologique de la société.

L'Open Data en tant que manifestation concrète du partage de l'information, peut ainsi être interprétée comme la mise en pratique des théories développées à la fin du XXe siècle par les recherches sur l'innovation, et qui répondent avant tout à des préoccupations économiques⁷. David et Foray (2003) listent ainsi les principaux arguments en faveur du partage au nom de l'innovation : l'échange d'information rompt les carcans des formes conventionnelles de production des connaissances en favorisant la créativité individuelle ; en facilitant les interactions entre non-spécialistes au-delà des cercles d'initiés (chercheurs, ingénieurs) et grâce au développement de systèmes de gestion décentralisés de collecte de données, de calcul et de partage des résultats, il facilite de nouvelles formes d'apprentissage collectif⁸ ; en ouvrant les bases de données, le partage permet un meilleur contrôle social de leur qualité, en multipliant les usagers susceptibles d'en corriger les défauts et d'identifier des problèmes⁹. L'un des processus vertueux supposément enclenché par le partage des données et fréquemment avancé comme argument central réside dans le changement d'échelle dans les capacités cognitives : ouvrir ses données, c'est permettre leur connexion (interopérabilité) à des ensembles plus vastes de données, donc d'opérer un changement dans les niveaux du savoir, que ce niveau soit appréhendé comme une complexité majeure, ou comme un espace plus grand. Ce saut d'échelle est ainsi supposé ouvrir de nouvelles possibilités d'actions auparavant interdites par la fragmentation thématique ou spatiale de l'information. Enfin, le partage en ligne est fréquemment justifié comme un moyen de limiter la redondance des études et recherches financées par les pouvoirs publics en limitant la perte de mémoire collective par la création de vastes bases ouvertes d'archivage, et donc *in fine* d'optimiser les dépenses.

Cette conception instrumentale de la diffusion de l'information environnementale constitue souvent le substrat de fond des textes et accords

⁷ La généalogie des théories et des débats sur les déséconomies sociales dues aux freins à la circulation de l'information est cependant complexe, et ce thème est discuté bien avant les années 1990 (voir entre autre : Russo, 1966).

⁸ Dans leurs travaux sur les communautés scientifiques, les logiciels libres ou encore les communautés d'utilisateurs plusieurs publications du MIT (Chesbrough 2003, Von Hippel 2002, Lakhani et Von Hippel 2001, Riggs et Von Hippel 1994) théorisent la notion d'innovation distribuée ou innovation ouverte basée sur le partage, la collaboration et la sérendipité et prônent le développement de systèmes favorisant des « *expérimentations multiples et décentralisées* » plutôt que les systèmes hiérarchisés dans lesquels la fonction d'expérimentation est accaparée par une unité centrale.

⁹ Au début des années 2000, ces postulats s'appuient sur l'affirmation du « coût social » que suppose la fermeture généralisée des bases de données en termes de perte de capacité d'innovation, notamment par la perte de capacité sociale à exploiter le potentiel de Gisements d'information (David et Foray donnent pour exemple le frein aux découvertes scientifiques que constitue l'impossibilité à mettre en relation des bases de données médicales dispersées dans le monde).

internationaux qui la promeuvent, reléguant selon nous à un second plan les considérations d'ordre démocratique, pourtant systématiquement présentées comme la principale explication à l'essor du partage public de l'information, selon laquelle l'Open Data serait le produit naturel d'une demande sociale croissante d'accès à l'information. Mais même le fameux « principe 10 » de la Déclaration de Rio sur l'environnement et le développement, couramment utilisé comme référence par les mouvements sociaux luttant pour la concrétisation du triptyque « droit à l'information - droit à la participation - droit à la justice en matière environnementale », justifie d'abord l'accès à l'information au nom de l'optimisation de la gestion, avant que d'en faire une question de droit¹⁰. Cela dit, c'est probablement l'interaction entre cette conception instrumentale et les nouvelles exigences démocratiques qui est à l'origine d'un vaste mouvement de reconnaissance officielle de l'information environnementale au rang de droit, traduite en conventions internationales (Aarhus en 1998), directives européennes (Inspire en 2007), et législations nationales¹¹. Une fois établis des dispositifs légaux de reconnaissance du droit à l'information environnementale et/ou de devoir de l'administration d'orchestrer son partage, ceux-ci peuvent être saisis par divers mouvements sociaux afin d'exiger l'application de ce droit. Les conflits environnementaux par exemple, et leur judiciarisation croissante (Melé, 2011), sont ainsi un puissant vecteur d'actualisation de cette nouvelle génération de droit et d'activation des dispositifs légaux en la matière. Selon cette seconde façon de penser le développement de l'Open Data, les diverses fortunes des initiatives de partage dépendraient avant tout de négociations difficiles entre des acteurs exigeant un meilleur partage de l'information publique (la « société civile ») et divers acteurs leur opposant de multiples inerties d'ordre bureaucratique ou politique (l'Etat). Cette explication semble néanmoins limitée et porteuse d'une vision simpliste de l'Etat, tant sont évidentes sur ce sujet les tensions au sein même de cet « Etat », entre expressions volontaristes et blocages institutionnels. Les difficultés que connaît la puissance publique pour mettre en place une politique d'Open Data dans un pays comme la France¹² semblent ainsi relever plus d'une absence de consensus entre les composantes de ce même Etat sur le sens à donner à une telle initiative, que d'un refus systématique de partager les données publiques.

On peut cependant avancer une autre explication au développement actuel des formes de partage des données publiques, non-concurrente des deux précédentes,

¹⁰ “Environmental issues are best handled with the participation of all concerned citizens, at the relevant level. At the national level, each individual shall have appropriate access to information concerning the environment that is held by public authorities, ...and the opportunity to participate in decision-making processes. States shall facilitate and encourage public awareness and participation by making information widely available” (nous soulignons).

¹¹ Pour un bilan des législations nationales en la matière en Amérique latine, voir : ECLAC, 2013.

¹² A titre d'exemple, le récent rapport intitulé « Ouverture des données publiques : les exceptions au principe de gratuité sont-elles toujours légitimes ? » remis au Premier Ministre le 5 novembre 2013 par M. Trojette, magistrat à la Cour des Comptes, pose clairement la question du maintien des redevances perçues par l'administration pour la vente de données publiques payantes.

en postulant que l'importance actuelle de l'information est tellement bien établie dans les dispositifs actuels de gouvernance qu'elle impose de nouvelles contraintes à l'Etat¹³, ces contraintes étant à l'origine d'un auto-renforcement des initiatives de partage. La légitimité de l'Etat dans le champ environnemental ne dépend désormais plus uniquement de ses prérogatives normatives et contraignantes, mais également de sa capacité à réguler la circulation de l'information environnementale, à la qualifier et à en organiser les flux. La multiplication des acteurs et de leurs capacités à produire et discuter l'information environnementale, notamment grâce à la diffusion des outils numériques, a placé l'Etat au sens large dans une situation de concurrence croissante dans le domaine de l'information. Ses actions peuvent de plus en plus être contestées par des acteurs non publics produisant et diffusant des connaissances environnementales, sur Internet, par exemple avec le développement de la cartographie 2.0 (Noucher et Gautreau 2013 ; Joliveau et al., 2013). Cette nouvelle concurrence serait ainsi à l'origine de certains développements de l'Open Data, l'Etat cherchant à maintenir une place centrale dans la formulation des questions environnementales en se repositionnant, entre autres lieux, sur le Web (voir dans ce numéro l'article de Gautreau et al.).

1.2. Le partage des données est-il vraiment un nouvel « instrument » des politiques publiques d'environnement ?

Replacer les initiatives de partage de données par les pouvoirs publics dans l'ensemble des instruments de la gouvernance de l'environnement constitue une seconde difficulté pour comprendre le phénomène. Le travail de définition et d'analyse des « instruments d'action publique » (Lascoumes et Le Galès, 2004b) par les sciences politiques, la sociologie de l'Etat et celle des techniques offre un cadre intéressant pour y resituer la mouvance Open Data et souligner ses particularités. Certes, celle-ci semble clairement relever des instruments « informatifs et communicationnels », et elle est présentée par ses tenants au sein de l'administration comme une véritable politique de l'Etat, ayant vocation à fournir un accès universel à l'ensemble des informations que celui-ci possède, ou à constituer un service public de prospection dans des « gisements » de données. L'Open Data est ainsi présenté comme une extension aux citoyens et à l'initiative privée de systèmes d'information

¹³ Dans un domaine initialement sans lien avec l'Open Data, Waterton et Wyne (2004) ont montré combien la façon dont on produit et diffuse l'information institutionnelle pouvait constituer désormais un élément fondamental de légitimité au sein d'instances publiques telles que les composantes de l'administration de l'Union Européenne. Leur étude détaille comment l'Agence Environnementale Européenne a lutté depuis sa création à la fin des années 1980 pour son indépendance face à la Direction de l'Environnement de la Commission, en mettant notamment en avant ses capacités à produire une information prétendument politiquement « neutre » et « objective ». C'est en se posant en garante d'un certain « modèle d'information » qu'elle a pu défendre l'originalité et l'indépendance de ses missions, dans le but d'éviter de devenir un simple outil de fourniture d'informations au service de l'administration.

autrefois restreints aux enceintes publiques, au service d'un projet plus vaste de pilotage de notre environnement grâce au partage de l'information.

Cependant, plusieurs caractéristiques de l'Open Data l'écartent de ces « instruments d'action publique » tels que définis par Lascoumes et Le Galès (2004b, p 13), c'est-à-dire « constitu[ant] un dispositif à la fois technique et social qui organise des rapports sociaux spécifiques entre la puissance publique et ses destinataires en fonction des représentations et des significations dont il est porteur ». La nécessité de l'Open Data est souvent plus justifiée au nom de valeurs (la transparence, la justice...) qu'au nom d'objectifs opérationnels spécifiques, démontrant par là même les multiples tâtonnements de l'Etat quant au rôle à assigner à ces pratiques disparates qui sont loin de former un tout cohérent. La question de la « certification »¹⁴ ou de la qualification des données partagées, c'est-à-dire le fait de les accompagner d'une description par des métadonnées (auteur, date de création, etc) constitue ainsi une ligne de partage entre deux grandes catégories d'initiatives. Certaines ne font de cette question qu'un point secondaire ou négligeable, notamment lorsqu'elles ont pour principal but de faciliter l'accès à un patrimoine de données dispersées et produites par des acteurs multiples¹⁵. Dans d'autres cas, l'ouverture des données doit être précédée d'un lourd travail de qualification de la donnée, qui constitue une responsabilité centrale de l'administration qui en a la charge¹⁶.

Plus largement, il faut s'interroger sur la capacité des initiatives Open Data à être porteuses d'une « organisation de rapports sociaux spécifiques entre la puissance publique et ses destinataires », et donc à être, plus qu'une simple technique de gouvernance, un véritable instrument, une « institution ». La promotion officielle de l'Open Data postule justement tout le contraire. Défendant l'idée qu'en diffusant des informations accessibles à tous, il n'existe plus –théoriquement– de destinataire précis de cette diffusion, donc aucune définition d'un « rapport social spécifique ». Dans ce

¹⁴ “Clearly, new methods need to be devised to ‘certify’ the knowledge circulating on the Internet within a context where inputs are no longer subject to control (unlike the knowledge disseminated by scientific journals, for example, whose quality and reliability are validated through the peer review process)” (David et Foray 2003).

¹⁵ C'est par exemple le cas du GBIF, base de données internationale sur les occurrences d'espèces animales et végétales financée par les membres de l'OCDE (<http://www.gbif.org/>), ou de certaines Infrastructures de données Géographiques donnant accès à des couches d'information géoréférencées. Ces initiatives proposent un accès à des données sur la qualité desquelles elles ne se prononcent pas directement, se contentant d'exiger de leurs auteurs qu'ils décrivent de façon standard la nature de ces données (fiches de métadonnées). Le très grand nombre d'auteurs parmi les données mises en ligne sur ces sites interdit en effet à l'administration qui les gère de vérifier au cas par cas leur mode de production.

¹⁶ C'est le cas de la diffusion de référentiels environnementaux, tels que les périmètres réglementaires (aires protégées, etc), les listes officielles d'espèces ou leurs aires de répartition par le Muséum National d'Histoire Naturelle en France (<http://inpn.mnhn.fr>), ou les cartes de végétation du Brésil par le Ministère de l'environnement (<http://mapas.mma.gov.br/mapas/aplic/probio/datadownload.htm>). Les administrations en charge de leur diffusion se portent garante de la qualité attribuée aux données, notamment en fournissant aux usagers des métadonnées sur leur constitution.

même ordre d'idées, la faible coordination entre les initiatives publiques de partage des données et la nature extrêmement hétérogène des formats, des dispositifs, des thèmes, interdit de penser que l'Open Data soit actuellement porteur d'une vision globale de l'Etat sur la façon de gérer la chose publique, dont l'environnement. Ainsi, l'Open data serait avant tout un ensemble de « techniques », propres aux différents acteurs publics, qui les exercent rarement de façon convergentes.

2. CE QUE CHANGE LE PARTAGE EN LIGNE DES DONNEES ENVIRONNEMENTALES : QUELQUES DEBATS SUR OPEN DATA ET EFFETS D'INFORMATION

Comprendre les effets sociétaux du partage de l'information constitue un défi important des recherches en environnement, tant les méthodes et les théories pour aborder cette question sont encore loin d'être stabilisées. Les « effets d'information » en matière environnementale sont encore très difficiles à identifier, sans que l'on sache toujours si cela est plus dû à l'inadaptation des dispositifs de recherche qu'au poids toujours écrasant d'autres facteurs plus décisifs (Gerstlé, 2010). Les recherches sur les usages des dispositifs de partage de données en ligne constituent une piste intéressante, encore limitées à des cas d'étude et peu systématisées¹⁷, et c'est à une généralisation de cette approche par les usages qu'appelle dans ce numéro l'article de Matthieu Noucher. En dépassant le paradigme communicationnel pour ne pas limiter l'Open Data à un simple vecteur d'information, dont il suffit d'étudier en amont l'encodage et en aval le décodage, une entrée par les usages doit permettre de prendre la mesure de la diversité des fonctions et rôles remplis par les données environnementales. La « donnée » est souvent cantonnée à un rôle de représentations graphiques, cartographiques ou statistiques du territoire ; comme le propose Florence Jacquinod (2014) pour les géovisualisations 3D, en ayant recours aux théories de l'acteur réseau il s'agirait alors d'examiner ses autres fonctions (expliquer, convaincre, discuter).

Une autre piste pour approcher certains changements induits par la mise en ligne de données environnementales consiste à recenser les principaux débats que celle-ci suscite, pour mettre en évidence les enjeux identifiés par les acteurs qui participent à ces initiatives.

¹⁷ A titre d'exemples, voir : Feyt et Noucher, 2014 sur les conséquences de l'accessibilité de nouveaux jeux de données dans la définition des référentiels géographiques ; Georis-Creuseveau, 2013 à propos de la contribution des infrastructures de données géographiques à la gestion des territoires côtiers ; Gautreau et Vélez, 2011 à propos de l'impact de la disponibilité de données en ligne sur les formes de résolution de conflits environnementaux.

2.1. Ouvrir ses données... et après ? La récurrente question des « usages » des données mises en partage

En 2012, les coordinateurs de l'une des initiatives mettant en ligne le plus grand nombre de données de biodiversité au monde –celle du GBIF (voir la fiche d'information dans ce même numéro) et ses près de 400 millions d'occurrences d'espèces animales et végétales- ont lancé une série de consultations auprès des groupes qui animent leur réseau, afin d'améliorer la visibilité et l'usage de leurs données. Organisé autour d'une plateforme en ligne à laquelle des organismes ayant passé une convention avec le consortium peuvent « connecter » leurs listes d'occurrences d'espèces, le GBIF s'articule autour de points nodaux nationaux qui coordonnent les actions au niveau de chaque pays partenaire. Après quelques années de fonctionnement, c'est le constat que le site était très faiblement cité dans la littérature scientifique qui a provoqué chez les coordinateurs le besoin d'imaginer les façons de rendre la plateforme plus attractive pour la communauté des scientifiques, essentiellement des naturalistes. Cet exemple est emblématique de plusieurs tensions qui traversent actuellement l'Open Data environnementale, qui à certains égards témoignent d'une maturation du processus. Après la mise en ligne massive de données, la question des usagers ou destinataires de ces initiatives, initialement abordée marginalement, revient au devant des préoccupations. Le mythe d'initiatives destinées à « tous » tombe, d'autant plus en contexte de restrictions budgétaires où leurs financeurs commencent à vouloir évaluer le retour sur investissement de ces expériences. La volonté du GBIF de se recentrer plus clairement sur la communauté naturaliste s'est accompagnée d'une évolution des fonctionnalités du site web, afin d'offrir de nouveaux services permettant notamment de mieux évaluer la structuration et la qualité des jeux de données téléchargés, une demande particulièrement forte de la part des usagers zoologues, biologistes ou écologues.

D'autre part, les discussions internes au GBIF ont permis d'identifier certains facteurs expliquant ce faible recours à la plateforme de la part des naturalistes : les données portent pour l'essentiel sur l'occurrence d'espèces, et non sur des propriétés fonctionnelles ou écosystémiques du vivant, aujourd'hui essentielles pour la compréhension des dynamiques de la biodiversité. A ainsi été mise en évidence l'une des faiblesses inhérentes aux initiatives Open Data de cette ampleur : rechercher une portée internationale suppose, pour ce type de base ouverte, de cibler en priorité les données facilement harmonisables et les plus abondantes, et de laisser en second plan des données plus complexes, mais plus difficilement harmonisables¹⁸. Bien que ce soit en ciblant justement des données « simples » que ces initiatives ont réussi à se développer, on voit par cet exemple que le caractère souvent unidimensionnel de leurs contenus constitue rapidement un frein à l'essor de leur usage. Les difficultés auxquelles se heurte actuellement le Système d'Information sur la Nature et les

¹⁸ Dans le cas des données de biodiversité, les données génétiques et paysagères sont plus difficiles à harmoniser entre elles, tant les protocoles de collecte diffèrent.

Paysages¹⁹ en France pour incorporer des données sur le paysage constituent un autre témoignage des dilemmes auxquels sont soumises ces initiatives : discuter de manière approfondie des méthodes pour créer d'emblée des bases ouvertes complexes pouvant intéresser une grande diversité d'utilisateurs, mais au prix d'un temps très long de mise en place, ou mettre en place rapidement des bases ouvertes, mais au prix d'une certaine pauvreté thématique de celles-ci, et d'une réduction du nombre potentiel d'utilisateurs intéressés²⁰.

2.2. L'Open Data est-il un facteur de réduction de la diversité des savoirs sur l'environnement ? Normalisation et filtrages amont-aval des connaissances

En lien direct avec les « dilemmes » évoqués au paragraphe précédent, un autre débat de fond sur l'Open Data se réfère à ses potentiels effets négatifs sur la diversité des modes de connaissance de l'environnement. L'injonction du partage des données s'accompagne aujourd'hui inévitablement de discussions sur les outils informatiques de cette mise en partage, et donc sur le reformatage de données produites par des acteurs différents et selon des méthodes différentes, afin qu'elles puissent être agrégées, comparées, combinées dans des systèmes d'information englobants. Nous venons de voir au paragraphe précédent comment certaines initiatives ouvertes jouent un rôle de « filtre aval » de la diversité des connaissances, en mettant prioritairement en ligne les données les plus facilement combinables, et laissant de côté les données trop complexes ou hétérogènes. Certains auteurs estiment que les processus de partage, lorsqu'ils sont fortement pilotés par la puissance publique, peuvent également avoir un effet de « filtrage amont » de la diversité des connaissances, impactant sur les formes mêmes de construction de ces connaissances.

Dans leurs travaux portant sur la mise en place du Système d'Information sur la Nature et les Paysages et sur la « Trame verte et bleue » française, Pierre Alphanhéry et Agnès Fortier signalent le risque d'un recul des formes locales de connaissance naturaliste (Alphanhéry et Fortier, 2011; Alphanhéry et al., 2012b) : le besoin de normalisation des données pour les intégrer aux systèmes partagés suppose l'élaboration de protocoles qui impacteraient sur l'amont des processus de collecte

¹⁹ Le SINP (<http://www.naturefrance.fr/sinp/presentation-du-sinp>) a vocation à faciliter le partage des données sur la biodiversité et les paysages français. Il se structure autour d'actions de mise en réseau des acteurs naturalistes, de diffusion d'outils informatiques de saisie des données, et de processus de normalisation et de standardisation de ces données (définitions, établissement de protocoles de saisie et de traitement...). En 2013, le volet « paysage » a été de facto suspendu, notamment par manque de consensus sur le type de données pouvant être intégrées au système pour rendre compte de cet objet complexe et de ses évolutions. Les progrès du SINP sont beaucoup plus nets dans le domaine des données de biodiversité.

²⁰ Les dilemmes de la constitution de bases de grande ampleur spatiale ou portant sur des objets complexes (ce qui est toujours le cas des objets environnementaux) sont nombreux, que les bases soient ouvertes ou non. Waterton & Wyne (2004) rapportent à propos des premières bases de données de l'Agence Environnementale Européenne que la décision de n'y incorporer que des données existantes pour tous les pays de l'Union a longtemps interdit d'intégrer les données sur le nucléaire, pourtant fondamentales pour la gestion des risques.

(mode d'observation naturaliste) et donc sur la façon même dont les acteurs locaux créent leurs connaissances, avant de les verser dans des systèmes ouverts à échelle régionale ou nationale²¹. Les dynamiques enclenchées au nom du partage par la coordination nationale d'acteurs locaux sous l'égide du Muséum National d'Histoire Naturelle auraient pour autre effet de « professionnaliser » les acteurs associatifs, induisant *in fine* une homogénéisation et une normalisation de leurs pratiques de collecte des données. Si la portée réelle des pertes de diversité des savoirs naturalistes liées à ces processus reste selon nous à mieux démontrer, le débat soulevé a le grand mérite de signaler les potentiels dangers de processus qui sont loin de se résumer à la mise en ligne de données publiques : le partage des données en général, et l'Open Data en particulier supposent de profonds réagencements institutionnels qui impactent notre façon de mesurer, de concevoir et d'appréhender notre environnement.

3. LES APPORTS D'UNE APPROCHE GEOGRAPHIQUE DE L'OPEN DATA ENVIRONNEMENTALE

Investie prioritairement par les sciences politiques ou la sociologie, la question de l'Open Data se prête pourtant bien aux approches géographiques, qui peuvent en éclairer de façon originale certaines de ses dimensions, ce dont ce numéro thématique rend en partie compte. La nature fréquemment géographique des données des systèmes d'information faisant l'objet d'un partage en ligne exige que soient développées des méthodes de formalisation du contenu de ces systèmes, afin de mieux comprendre la portée spatiale des initiatives. Les recompositions des rapports entre acteurs agissant à des niveaux territoriaux différents en lien avec des initiatives de partage des données supposent là aussi des analyses géographiques. Enfin, il est nécessaire de différencier les modalités de mise en place d'initiatives Open Data et leurs effets potentiels selon les contextes géographiques, notamment selon les trajectoires nationales de développement.

3.1. Une « political ecology » de l'Open Data

Les recherches en *Political Ecology*, fortement influencées par la géographie, portent notamment sur le lien entre formes d'appropriation de l'espace, genèse des savoirs sur l'environnement, et rapports de pouvoirs. Elles sont appropriées pour replacer le phénomène Open Data dans le temps long de la construction des savoirs environnementaux, et pour mieux différencier ses enjeux selon les territoires, notamment les pays émergents et en développement. Il semble ainsi nécessaire de relativiser en fonction des contextes nationaux les préventions de certaines approches foucaaldiennes des systèmes d'information, qui y voient avant tout des dispositifs de

²¹ En France, une part majeure des données de biodiversité (occurrences d'espèces) est collectée par des associations d'amateurs, qui les fournissent selon diverses modalités aux autorités nationales (Muséum National d'Histoire Naturelle).

contrôle social par le biais de la normalisation des connaissances et par leur centralisation. Dans un grand nombre de pays en développement, l'absence de bases de données cohérentes, normalisées, centralisées sur l'ensemble du territoire (et *a fortiori* de données partagées en ligne) constitue un frein au traitement démocratique des questions environnementales et au développement d'une action de l'Etat répondant à des principes de justice environnementale, notamment en termes d'exposition au risque. Plus simplement encore, l'absence de données dans un grand nombre de pays rend tout à fait illusoire toute volonté de gouvernance informationnelle de l'environnement²². Dans un registre proche, l'article de Louca Lerch sur l'Infrastructure Nationale de Données Géographiques bolivienne montre que les initiatives Open Data peuvent aussi constituer des outils de reconstitution d'une capacité informationnelle, par des Etats ayant pendant longtemps perdu leur capacité à produire et certifier des connaissances nationales sur leur territoire et leur environnement. Dénoncés dans certains territoires tels que la France comme un outil de recentralisation rampante de la production de connaissances²³, les initiatives Open Data peuvent donc dans d'autres contextes être mises au service de projets de reconquête de souveraineté de l'Etat.

En lien avec cette relativisation géographique des enjeux de l'Open Data, les approches politiques et historiques de la construction de capacités cognitives par l'Etat montrent également que les tentatives permanentes de constituer des connaissances homogènes sur la société sont très souvent marquées par l'échec. L'ouvrage de Scott « Seeing like a State » (1998) décrit la tendance historique de l'Etat à vouloir rendre « lisible » territoires et environnement par la mise en place de systèmes d'information et la normalisation des connaissances statistiques et cartographiques (cadastres). Dans cet ouvrage trop souvent lu avec un regard foucaldien simpliste y voyant uniquement l'avancée de techniques de contrôle social, Scott signale au contraire l'échec très fréquent de ces tentatives. L'Open Data, qui peut être assimilé à une ramification contemporaine de cette volonté de constituer des systèmes d'information sans cesse plus complexes sur la société, est également soumis à ce risque. La fragmentation et le caractère très limité des expériences actuelles ne permet pas encore de prouver que les dangers dénoncés par certains en termes de normalisation des savoirs et de standardisation des formes de concevoir et de gérer l'environnement soient réellement en germe dans ce vaste processus.

²² C'est la principale question que pose Mol (2009) à propos de la Chine et du Vietnam : "...what is the relevance of these new modes of informational environmental governance in the so-called informational peripheries where environmental information is known not to be made widely or freely available?"

²³ Cette critique est sous-jacente aux travaux précédemment cités de Pierre Alphanéry et Agnès Fortier.

3.2. Une géographie de l'information environnementale pour mieux comprendre l'Open Data

C'est également en termes méthodologiques que la géographie peut apporter une meilleure compréhension de l'Open Data, et tout d'abord en appliquant aux corpus de données mis en ligne des méthodes d'analyse spatiale déjà bien éprouvées dans d'autres champs. La représentation cartographique de ces corpus, dont l'article de Louca Lerch donne un aperçu dans ce numéro à propos de la Bolivie, permet de saisir la couverture spatiale et temporelle de certains territoires par des initiatives de partage. Dans le prolongement des études géohistoriques sur les progrès de la connaissance territoriale et environnementale par l'Etat moderne (extension des cadastres, développement de la cartographie topographique...), diverses méthodes de formalisation des corpus de données mises en ligne doivent être développées, tant pour comprendre les inégalités spatiales de représentation de certaines thématiques environnementales sur le Web (Gautreau, accepté) que pour mesurer les progrès inégaux de normalisation et standardisation de ces données selon les pays²⁴.

Des méthodes spécifiquement géographiques sont par ailleurs à inventer pour contribuer à la compréhension des processus de décentralisation ou au contraire de recentralisation de la gestion de l'environnement liées à la circulation des données. L'analyse de réseaux d'acteurs institutionnels peut en effet s'appuyer sur une étude des flux d'information entre ces acteurs, notamment par l'exploration des réseaux de leurs sites web (par exemple, réseau formé par les sites des DREAL des régions françaises et les sites ministériels et scientifiques comme celui du Muséum National d'Histoire Naturel). L'étude de la circulation de certains jeux de données dans ces réseaux permet d'aborder des rapports de hiérarchie et de coordination/absence de coordination entre ces acteurs.

La géographie offre enfin la possibilité de remobiliser les apports des courants relevant de la cartographie critique (Harley, 1989) pour déconstruire certains discours sur la portée ou la nature des données cartographiques en accès libre. Laurent Couderchet et Xavier Amelot (2010) offrent un exemple de cette démarche, en démontrant à partir d'une analyse spatiale fine du fichier de l'inventaire naturaliste des ZNIEFF²⁵ mis à disposition sur le web comporte d'importantes hétérogénéités²⁶, liées au mode décentralisé et fortement local du choix de ces zones, essentiellement par des acteurs associatifs naturalistes. Cette analyse géographique et critique déconstruit l'idée que ce fichier constitue un document représentant une vision homogène des zones

²⁴ Voir notamment une première ébauche d'analyse spatiale et temporelle des patrimoines de données contenus dans les Infrastructures Nationales de Données Géographiques d'Argentine, de Bolivie et du Brésil (<http://baguala.hypotheses.org/324>).

²⁵ L'inventaire ZNIEFF (pour Zones d'Intérêt Ecologique, Faunistique et Floristique) a été lancé en 1982. Il a pour objectif de recenser et de décrire des secteurs présentant de « fortes capacités biologiques et un bon état de conservation » (Site du Muséum National d'Histoire Naturelle).

²⁶ Sur la précision des contours des zones d'inventaire et sur les patterns géographiques de choix des zones.

d'intérêt pour la conservation, qui est pourtant à la base de la légitimité de son maintien et de son usage à diverses fins. La diffusion de bases de données géographiques réutilisables et devenues malléables incite d'ailleurs la géographie à envisager un renouvellement partiel de ces approches critiques (Joliveau et al., 2013). Hanna et Casino (2003) en appellent ainsi à ne plus considérer les cartes comme de simples représentations de lieux mais comme « des sujets mobiles et pleins de sens à travers un ensemble de pratiques socio-spatiales conflictuelles, complexes, intertextuelles et interreliées. »

L'Open Data, tel que le soulignent Samuel Goëta et Jérôme Denis, se construit sur le mythe que les données « libérées » sont universellement réutilisables et peuvent faire l'objet d'une multitude de nouveaux usages²⁷. Par une approche ethnographiques, ils montrent comment les acteurs décidant de mettre en ligne leur données procèdent à un travail de « désindexicalisation » et de « délocalisation » de leurs données afin de leur donner cette apparence d'universalité, en bref, comment ils tentent de gommer les motifs initiaux de création de ces données et les traitements intermédiaires dont elles ont fait l'objet avant de pouvoir être mises en ligne. Mais ce travail de gommage reste en grande partie une entreprise de mystification, qui tente de masquer le fait que les jeux de données mis en ligne sont justement marqués par toute leur histoire précédant leur « libération ». C'est à un tel travail de déconstruction des pratiques de la mouvance Open Data qu'une cartographie critique renouvelée pourrait contribuer, particulièrement à propos des données géographiques ou de toute information géolocalisée, ce qui est le cas de la majorité des données environnementales.

BIBLIOGRAPHIE

- ALPHANDERY, P., DJAMA, A., FORTIER, A., FOUILLEUX, E. (Eds.) (2012a), *Normaliser au nom du développement durable*, Quae. ed. Versailles, France.
- ALPHANDERY, P., FORTIER, A. (2011), Les associations dans le processus de rationalisation des données naturalistes, *Natures Sciences Sociétés*, 22–30.
- ALPHANDERY, P., FORTIER, A. (2012), Normes, savoirs et pouvoirs dans la production des données naturalistes en France, in: ALPHANDERY, P., DJAMA, A., FORTIER, A., FOUILLEUX, E. (Eds.), *Normaliser au nom du développement durable*, Quae, Versailles, France, pp. 141–156.
- ALPHANDERY, P., FORTIER, A., SOURDRIL, A. (2012b), Les données entre normalisation et territoire : la construction de la trame verte et bleue. *Développement durable et territoires*, 3 (2), [en ligne].
- CHESBROUGH H. (2003), *Open Innovation, the New Imperative for Creating and Profiting from Technology*, Harvard Business School Press.

²⁷<http://coulisses-opendata.com/2013/02/27/ouvrir-la-boite-noire-de-lopen-data-quelques-premieres-pistes-issues-des-coulisses/>

- COUDERCHET, L., AMELOT, X., (2010), Faut-il brûler les Znieff? *Cybergeog: European Journal of Geography*, [en ligne].
- DAVID, P.A., FORAY, D., (2003), Economic Fundamentals of the Knowledge Society, *Policy Futures in Education*, 1, 20–49.
- ECLAC, U.N. (2013), *Access to information, participation and justice in environmental matters in latin america and the caribbean: situation, outlook and examples of good practice*, United Nations.
- FEYT G. et NOUCHER M. (2014), La gouvernance informationnelle, outil et enjeu stratégiques des recompositions territoriales : vers l'émergence de nouveaux référentiels géographiques ? *Colloque « Fronts et frontières des sciences du territoire »*, GIS CIST, 27-28 mars 2014, Paris.
- GAUTREAU, P. (accepté), Ce que change le partage en ligne des données environnementales. Information environnementale, internet et rapports de pouvoir dans trois pays sud-américains, in: *Colloque d'Orléans « Géographie, Écologie, Politique : Un Climat de Changement »*, 6-7 Octobre 2012.
- GAUTREAU, P., VELEZ, E. (2011), Strategies of environmental knowledge production facing land use changes: Insights from the Silvicultural Zoning Plan conflict in the Brazilian state of Rio Grande do Sul, *Cybergeog: European Journal of Geography*, [en ligne].
- GEORIS-CREUSEVEAU J., GOURMELON F., CLARAMUNT C., MAULPOIX A., MARECHAL C., BÖHM A. (2013), *Usages des infrastructures de données géographiques par les acteurs côtiers français*, Actes du Séminaire M@rsouin, St Malo.
- GIRAUD P.A. (2014 - à paraître), Appartenir à des lieux d'une mouvance : le cas du Libre, In SAULE-SORBE H. et PEYLET G. (eds), *L'appartenir en question*, Presses Universitaires de Bordeaux, coll. Eidolôn.
- GERSTLE, J. (2010), Effets d'information, in: *Dictionnaire Des Politiques Publiques*, 3e Édition Actualisée et Augmentée, pp. 218–225.
- HARLEY J.B. (1989), Deconstructing the map, *Cartographica*, 25 (2), 235-252.
- JACQUINOD F. (2014), *Production, pratique et usages des géovisualisations 3D dans l'aménagement du territoire*, Thèse de doctorat de l'Université de St Etienne, 604 p.
- JOLIVEAU T., NOUCHER M., ROCHE S. (2013), "La cartographie 2.0, vers une approche critique d'un nouveau régime cartographique", *L'Information Géographique*, 77 (4), 29-46.
- LAKHANI K. et VON HIPPEL, E. (2001), *How open source software works : « Free » user-to-user assistance*, MIT Sloan School of Management Working Paper, 4117.
- LASCOUMES, P., LE GALES, P. (2004a), L'action publique saisie par ses instruments., in: LASCOUMES, P., LE GALES, P. (Eds.), *Gouverner par les instruments*, Presses de Science Po, Paris, pp. 11–46.
- LASCOUMES, P., LE GALES, P. (Eds.) (2004b), *Gouverner par les instruments*, Presses de Science Po, Paris.

- LAVOUX, T. (2003), L'information environnementale: nouvel instrument de régulation politique en europe? *Revue Internationale de Politique Comparée*, 10, 177–194.
- MELE, P. (2011), *Transactions territoriales. Patrimoine, environnement et actions collectives au Mexique*, Presses Universitaires François Rabelais, ed. Tours.
- MOL, A.P.J., (2009), Environmental governance through information: China and Vietnam, *Singapore Journal of Tropical Geography*, 114–129.
- NOUCHER M, GAUTREAU P. (2013), Le libre accès rebat-il les cartes ? *Les Cahiers du numérique*, 9 (1), 57-83.
- RIGGS W., VON HIPPEL E. (1994), The impact of scientific and commercial values on the sources of scientific instrument innovation, *Research Policy*, 23, pp. 459-469.
- RUSSO S. (1966), L'information, source de pouvoir, *Economie rurale*, 69, 3–10.
- SCOTT, J.C. (1998), *Seeing like a state: how certain schemes to improve the human condition have failed*, Yale University Press, New Haven.
- VON HIPPEL E. (2002), *Open source projects as horizontal innovation networks – by and for users*, MIT Sloan School of Management Working paper, 4366-02.
- WATERTON C & WYNNE B. (2004), Knowledge and politics in the European Environmental Agency, in: Jasanoff S., *States of knowledge: the co-production of science and the social order*, Routledge, London, New York : 87-107.